Shared Decision-Making Forward an Autonomous Navigation for Intelligent Vehicles

Ahmad Shour, Hugo Pousseur, Alessandro Corrêa Victorino, Veronique
Cherfaoui

Shared Decision-Making Forward an Autonomous Navigation for Intelligent Vehicles*

Ahmad Shour[+], Hugo Pousseur[+], Alessandro Correa Victorino[+], and Veronique Cherfaoui[+]

Abstract— Shared control is a situation where humans and machines carry out a task in a cooperative way, and it is a popular approach to facilitate control and communication between humans and intelligent systems. Automated driving vehicles that share control with human drivers have been shown to acts better than fully automated systems and manual driving. However, there is little consensus in guidelines for the design and definition of shared control, some researchers assist humans using haptic feedback, and others use an optimization process and many other works. In this paper, Non-Cooperative Game Theory is applied for the Decision-Making to formulate the final input to the vehicle in the form of linear velocity and steering wheel angle, where a joystick is used for manipulating human inputs and a computer vision technology which is the "Visual Servoing" implemented to command the autonomous driving inputs, after giving a value for the blending coefficient which is determined by a function dependent of the admissibility of inputs and the similarity between them. The final navigation decision sent to the intelligent vehicle is the output from the fusion system. Noting that, the fusion system has the only right to command the driving inputs to the vehicle. Finally, simulation results are presented with the help of SCANeR Studio and MATLAB/SIMULINK Simulators.

I. INTRODUCTION

A. Motivation

Autonomous navigation of vehicles is an important topic nowadays. Most road accidents are caused by human errors [1], such as making a phone call, being drunk, or even when they are fatigued. Even the application of the most recent Advanced Diving Assistance Systems (ADAS) has helped to reduce accidents, the development of on-boarded autonomous navigation systems applied to intelligent vehicles has appeared as a potential solution. Automated vehicles can be the rescue of human lives that are dying each day while driving. Google has already started the testing of their autonomous car "Waymo self-driving cars" on the streets, also the organization of challenges for autonomous cars continues with DARPA (Defence Advanced Research Program of America). In this event, research centers and universities around the world participate and display their research and development in the domain of autonomous cars. Fully autonomous driving is planned through the five series of levels defined by the American SAE J3016 standard [2], currently, the focus is relaid on the development of the third and the fourth levels of autonomy.

In this paper, we present a methodology to solve the problem related to the fusion of the capabilities between two intelligent agents (Humans and Intelligent Systems) to drive a car. Keeping humans in the loop to interact with intelligent systems will lead to efficient results in the way of driving, in a shared control navigation strategy. In this strategy, we can avoid any part of weakness from both sides. Where, automated vehicles transcend the human driver in sudden situations, power, and more persistence over a long period. While humans are superior in cognition, reasoning, decision making, and more intelligent in the perception of the environment.

The shared control idea aims at synergistically merging the competencies of both humans and intelligent systems. Autonomous navigation for intelligent vehicles usually requires a control system for path following. This control system is composed of longitudinal and lateral controllers. Our novel approach in this paper is the "Shared Control" strategy to both longitudinal and lateral driving through cooperative autonomous vehicle navigation between human driver and autonomous driver.

B. Related Works

In [4], cooperative non-linear quasi-continuous sliding mode control in a steer-by-wire vehicle as a lane-keeping assistant is proposed, where the assistant system must give the authority to the driver when he/she is available. The focus was on the interaction between the human driver and the controller. This interaction was based on a variable parameter computed by the monitoring system of the driver. Then, the final steering shared angle between the human and controller is achieved depending on the human angle, automatic angle, and the computed parameter.

The strategy of the shared control can be dependable on the driver's intention. After planning the trajectory based on the kinematic safe envelope, driver's path error can be used to select the trajectory that the human wants to follow, where MPC used to implement path tracking [7].

Shared control for the longitudinal vehicle control [8], where both the human and the intelligent systems can apply a force on the accelerator pedal. An optimization problem is applied to optimize the interaction between them, through a game theory approach. Each driver is modeled as an objective function to follow the desired trajectory. Noting that, the intelligent systems can only apply an opposite force to the pedal, no braking force. This methodology was applied to

two scenarios, one for speed-following and the other for car-following, and efficient results were shown. Nash equilibrium between driver and intelligent system used to formulate a control algorithm for the longitudinal control of the vehicle.

In [3], another shared control methodology is presented between the user and intelligent assistance system via a haptic joystick force feedback. Focusing on the human intentions defined as the position error and the desired speed of the wheelchair, where unknown input observers are developed to estimate those intentions based on joystick motions. Tagaki-Sugeno fuzzy model provides the optimal fuzzy controller and LQR controller proposed to form the lateral and angular torque feedback to the joystick, respectively.

In [5], the experiments applied to show that haptic torque feedback on the steering wheel without any warning from the automated system is not comfortable for the human driver. Gestures can be used to warn the human before any action from the assistance system, for example, warning by sound what the automated system will act.

The methodology proposed in [9] is based on Shared control to stabilize an inverted pendulum taking inputs from humans and intelligent systems. The problem of the fusion between the two control inputs is modeled as the regulation of the conflict between them to zero. LQR controller is used, and the admissibility of each input is specified by the R and Q matrices of the controller. Noting that, the fusion system has the only right to give the final input to the cart to have a stabilized inverted pendulum. The role of the fusion system is to nullify the conflict between the two, that what called "Conflict Resolution".

C. Contribution

Till now, computer vision working alone on a task is still weak compared to working cooperatively with the human [6]. In this paper, a computer vision technology which is the "Visual Servoing" will be working cooperatively with the human to drive a vehicle. The proposed strategy is based on the Non-Cooperative Game Theory to formulate the final input for the vehicle, based on the fusion between the two independent inputs. The fusion of the two inputs is based on a mathematical strategy to choose the coefficient of blending. Fusion coefficient is defined by a function dependent on the admissibility of each input and the similarity between them. The methodology is validated using the professional driving simulator SCANeR Studio with the help of MATLAB/SIMULINK.

The rest of the paper is organized as follows. A background of the work is presented in II. The autonomous driving controller explained in III. The fusion methodology is computed in IV. Some experimental results are shown in V. Finally, conclusion is presented in VI.

II. BACKGROUND

A. Non-Cooperative Game Theory

Game Theory is a mathematical theory related to the study of decision-making in an interactive situation [12]. In this paper, the fusion problem is represented in terms of a game. Non-Cooperative Game Theory (NCGT) is applied to the human driver and the autonomous agent, which will present the two players in our work. Each player will provide the driving inputs from their perspective of the situation independent of the other. These two inputs are merged, in a fusion system formulated at the coming sections, to resulting a final input to the vehicle. The final driving inputs are considered as the longitudinal velocity and the steering wheel angle.

In [10], the methodology of the NCGT showed that the final scaled coefficient will be at the line between the two mapped points of the two agents, which is the Nash Equilibrium solution. The input data from the manual driving and autonomous driving of the vehicle will be fitted into a quadratic form: $y = ax^2 + bx + c$.

Then, these quadratic forms will be mapped into 2D points of coordinates (a,b). The coefficient c is not considered in the mapping because the output y is more sensitive to the coefficients a and b. For example, if we suppose that: $y_{human} = a_{human}x^2 + b_{human}x + c_{human}$; which is the fitted quadratic form of the human input, then the mapped point of this fitted data will be point H presenting it in 2D of coordinates (a_{human}, b_{human}).

The strategy of NCGT is related to maximizing the gain/payoffs. The final driving input computed by the final system can be different than the individual inputs. Which will be presented by the "Loss Utility Function". It represents the approximation of the driver resistance to the deviation from the intention. It is considered as a parabolic shape, and its function of the form:

$$f(x,y) = (x - x_1)^2 + (y - y_1)^2 \qquad (1)$$

where x_1 and y_1 are the coordinates of the mapped point of the driving input.

The functioning of the shared control system should be human-centered. In situations, where the human driving intention is very different than that of the autonomous system, the shared control system should become human-centric. With this functionality, the intelligent vehicle may also be termed as 'Human-Centered Autonomous Vehicle'.

The similarity between the two input profiles is inversely proportional to the euclidean distance between their respective mappings on the 2D plane. When the similarity is low then the intentions are different, we must follow the human in those situations. Since the NCGT is based on the loss utility function the driving intentions are compared in the context of their respective loss utilities, i.e. the loss utility functions of the respective mappings are compared to calculate the similarity measure. It has this form:

$$Sim(P_H, P_A) = \frac{1}{(1 + \delta * D^2)} \qquad (2)$$

where P_H and P_A are two profiles of the human and autonomous inputs, respectively. δ is the tuning parameter and D is the euclidean distance between the two mapped points.

Let A and H be the two mapped points of the autonomous and human speed inputs, respectively. As shown in (Fig. 1),

we present a simple example to show how the input data (speed inputs in this example) can be mapped into 2D points.
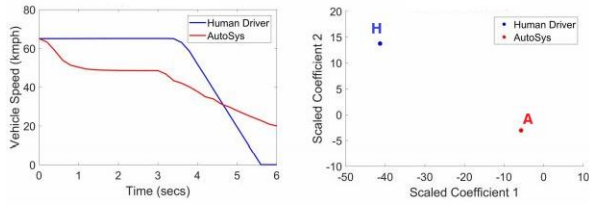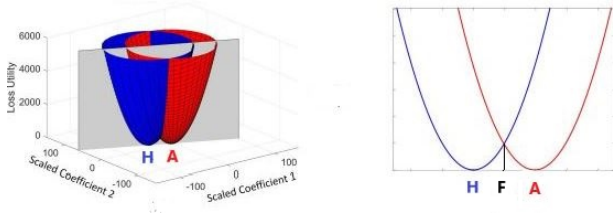

Fig. 1.   Speed inputs and there 2D mappings

Fig. 2.   Final point presentation

Then, for each input a loss utility function (Fig. 2) will be presented to be able to find point F, which is the bargaining solution between the human driver and the autonomous system. F belongs to the plane of intersection between the two parabolic shapes, that presents the loss utility of each input. As shown in (Fig. 2), in the 2D presentation that point F is the intersection between the two parabolas, this confirms that the F lies on the line between A and H.

B. Inputs Admissibility

Each parabolic function will be presented by an opening of its U-shape. Here, the definition of the opening comes from the principle of quality for each input.

Suppose that we the quality of the human inputs at a certain situation is higher, in another word, the human inputs are more admissible, then the opening of the parabola will be smaller than that for the autonomous parabola opening. Then the final input, in this case, will be more shifted to the human intentions.

That was presented by Belief Function Theory in our previous work [11], using the occupancy grid of the map with the help of lidar and camera sensor data to obtain the environment state. A degree of belief for each input in its decision admissibility "quality" is given a probability value.

III. VISUAL CONTROLLER

The visual controller drives the car in keeping the vehicle in the center of the current lane. The controller defines in divided into two parts: the first part detects lanes from the image of the front vehicle camera and the second part defines the command allowing to refocus the vehicle to the lane center.

A. Lane detection

A study shows, due to CNN, the lane detection accuracy increased from 80% to 90% compared with traditional image processing methods [14]. In lane detection, deep-learning compared to traditional detection methods can detect lanes in complex situations with shadows, occlusions, or curves [15]. The lane detection is divided into two parts, the first part creates binary masks, one per line from the raw image, and the second extracts a model for each line.

1) Masks prediction: The binary masks are predicted by two deep-learning models. These models are trained and tested on the CuLane dataset [16].

a) Mask prediction model: The first model, mask prediction, is an autoencoder that predicts lines mask of an instant t not depending on previous predictions from the camera image. The output of the model is four masks, each mask represents the mask activation for a line. The true positive are much less represented compared to the true negative, so the model has been trained in using the weighted binary cross-entropy loss function.

$$L(\hat{y}, y) = -(\hat{y} * log(y) * w_1 + (1 - \hat{y}) * log(1 - y) * w_0) \quad (3)$$

with $\hat{y}$ true value, y value predicted and w_0, w_1 weights for true negative/positive values.

b) Tracking model: The second model, tracking, is another autoencoder model where the coder part is defined by conv2D-lstm layers [17]. This model corrects prediction done by the first model based on previous predictions. This has been trained on noise data to learn to denoise data based on previous predictions. To improve the runtime execution, previous predictions are saved already encoded avoiding encoding previous predictions at each prediction by the tracking model.

As (Fig. 3) shows the prediction process from the raw image to line masks. As explained above, the tracking process improves the robustness of the prediction. This process can help to fill partial prediction, (Fig. 4), or remove outliers.

2) Lines regression: Once masks are predicted, a model needs to be defined to interpolate each mask line. This interpolation is performed into two steps, the first one applies RANSAC [18] algorithm to remove outliers and the second one defines the quadratic regression.

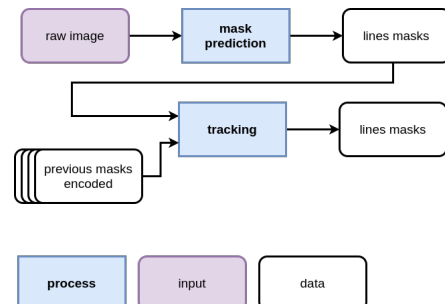As (Fig. 5) shows intermediate results from mask prediction until road segmentation.

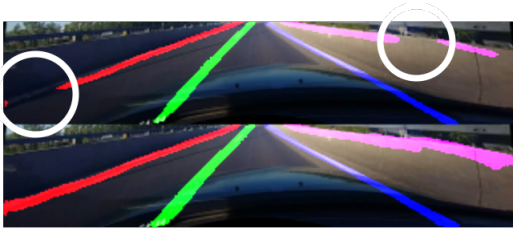Fig. 3.   Binary masks predictions process

Fig. 4. Tracking robustness example, the top image shows predictions without tracking and the bottom image shows predictions with tracking
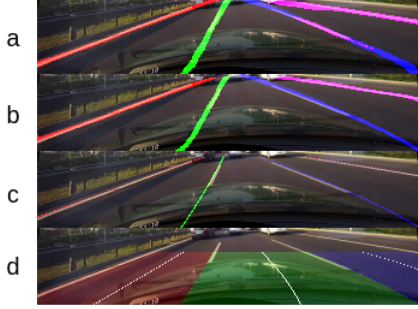

Fig. 5. Lane detection results, (a) masks prediction, (b) masks prediction with tracking, (c) regressions, (d) road segmentation.

B. Control Commands

Once lanes are detected, it's possible to control the define the command allowing to recenter the vehicle to the center of the vehicle. The solution is to apply the Visual Servoing method [19].

1) Visual Servoing: The visual servoing allows defining the vehicle commands from features image. In our case features image S are defined by the point $D = (X, Y)$, who represents the intersection of the path to follow the border of the image, and the angle Θ, who represents the angle between the tangent τ_P of the path P evaluated in D with the ordinate axis. Fig. 6 represents these features on an example. All vision-based control aims to minimize an error $e(t)$ defined as follow:

$$e(t) = S(t) - S^* \qquad (4)$$

With S features at the t instant and S^* desired values of these features. From this error, we can define row controller and column controller allowing us to reduce the error e.

Like denote [19], we can define a relation between features derivation ($\dot{S}$) and the vehicle velocity (v, w).

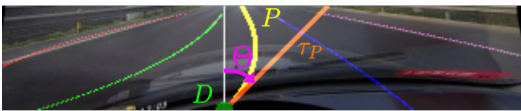$$\begin{bmatrix} \dot{X} \\ \dot{\Theta} \end{bmatrix} = A_{row}v + B_{row}w \qquad (5)$$

Fig. 6. Image features, point (D), path (P), path tangent (τ_P), angle between τ_P and $y - axis$ (Θ), others lines represents lane boundaries detected.

But $\dot{S}^* = 0$, so $\dot{e(t)} = \dot{S(t)}$ then if $B_{row} \neq 0$, the row controller law is defined by:

$$w = -B_{row}^+ (\begin{bmatrix} K_X e_X \\ K_\Theta e_\Theta \end{bmatrix} + A_{row}v) \qquad (6)$$

With K_X, K_Θ positive gains and B_{row}^+ the pseudo-inverse of B_{row}. In the same way, we can define the column controller law like this:

$$w = -B_{col}^+ (\begin{bmatrix} K_Y e_Y \\ K_\Theta e_\Theta \end{bmatrix} + A_{col}v) \qquad (7)$$

So from the current feature error and the linear velocity, we can define the angular velocity allowing to refocus the vehicle in the center of the current lane.

2) Image Based Dynamic Window Approach: The dynamic window approach (DWA) [20] is an online collision avoidance strategy, which means this strategy is based on the current states of the robot (i.e. current dynamics of the robot). The DWA method is divided into two steps: 1. define the limitation of the search space, 2. find the velocity maximizing the objective function. In our case, the first step is the same approach as described in [20]. The initial objective function is defined as follow:

$$G(v, w) = \lambda.heading(v, w) + \beta.dist(v, w) + \gamma.velocity(v, w) \qquad (8)$$

with the *heading* function measure the alignment of the target with the target, the *dist* function computes the *distance* to the closest obstacles and velocity function it's a simple project of the translation velocity.

The *heading* function can be replaced by another one based on visual servoing error [21]. Let denote $e_{t+\Delta t}$ like this:

$$e_{t+\Delta t} = \begin{bmatrix} X_{t+\Delta t} - X^* \\ Y_{t+\Delta t} - Y^* \\ \Theta_{t+\Delta t} - \Theta^* \end{bmatrix} = \begin{bmatrix} e_X \\ e_Y \\ e_\Theta \end{bmatrix} \qquad (9)$$

The *heading* function is modified as:

$$heading(v, w) = \lambda_1 XY_{error} + \lambda_2 \Theta_{error}(v, w) \qquad (10)$$

where:

$$XY_{error} = \begin{cases} 1 - \frac{|e_X|}{e_{Xmax}} & \text{for row controller} \\ 1 - \frac{|e_Y|}{e_{Ymax}} & \text{for column controller} \end{cases} \qquad (11)$$

$$\Theta_{error} = 1 - \frac{|e_\Theta|}{\pi} \qquad (12)$$

IV. FUSION METHODOLOGY

As shown in (Fig.7), the framework methodology of the "Fusion System" is presented. Where the human and the autonomous algorithm give their inputs independent of the other, each agent percept the environment in a different way. Here, the two parabolic drawings will not be formulated to find the final point and the final command. We will use mathematical functions to fuse the two inputs depending on the fusion coefficient will be discussed in the coming section. First, each agent will give his input independently to the other. Second, the input data will be fitted into quadratic form to plot the two mapped points to find the similarity between
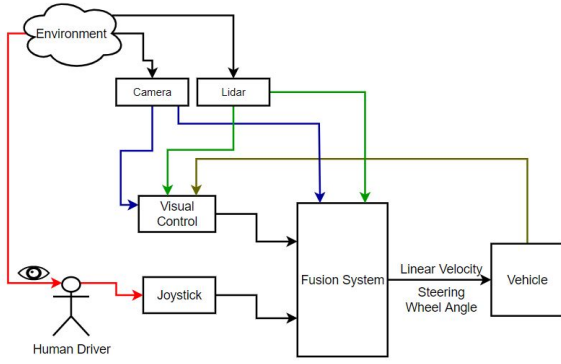
Fig. 7.   Fusion Framework

them. Third, a value of admissibility for each input will be given according to the methodology of Belief Function Theory. Finally, the coefficient of the fusion system will be computed using (Eq. 13), and the final input will be sent to the vehicle in the form of linear velocity and steering wheel angle.

A. Fusion Coefficient

The fusion coefficient Alpha "α" will be a combination between the admissibility and the similarity values (Eq. 13). Where α is between 0 and 1. As mentioned in II-B, the admissibility of each input is defined by a probability value between 0 and 1. Moreover, the similarity will be a boolean value, either 0 (the two inputs are non-similar) or 1 (the two inputs are similar).

$$\alpha = \Big((1-admh)\Big)*\Big((1-Sim)\Big) + \Big(\frac{max(0,admh-adma)}{admh-adma} + \frac{max(0,adma-admh)}{adma-admh}\Big) * \Big(Sim\Big) \quad (13)$$

$$\begin{cases} admh : admissibility\ of\ the\ human\ input \\ adma : admissibility\ of\ the\ autonomous\ algorithm\ input \\ Sim : Similarity\ between\ the\ inputs;\ 0\ or\ 1 \end{cases}$$

To explain, in a situation if both agents are making the same action, i.e stopping the vehicle then there is a similarity between them and the value of Sim=1 and the admissibility will be high since they are giving a correct decision. In another case, if the human decides with no reason to stop the vehicle, and the autonomous agent continues driving, in this situation there is no similarity between them, we should follow the human driver and Sim=0 even if the "adma" higher than "admh". A threshold limit will be defined to give the similarity "Sim" value.

B. Final Input

As mentioned before, F is the point of the bargaining solution, which means that it is the final point decision of the inputs and it lies between A and H in the 2D presentation of the mapped points from the quadratic fitting of the data.

The coordinates of F are the results of blending between the two mapped points of the human and autonomous system,

after specifying the admissibility of each input.

$$\begin{cases} x_F = x_H - \alpha(x_H - x_A) \\ y_F = y_H - \alpha(y_H - y_A) \end{cases} \quad (14)$$

In this way, point F will always lie on the line between the mapped points.

Finally, the final input to be sent to the vehicle is a fusion between the two inputs:

$$u_f = u_h - \alpha * (u_h - u_a) \quad (15)$$

$$where : \begin{cases} \alpha \in [0,1] \\ u_f\ is\ the\ final\ input \\ u_h\ is\ the\ human\ input \\ u_a\ is\ the\ autonomous\ input \end{cases}$$

This fusion will be applied for the linear velocity input and the steering wheel input.

C. Driving Modes Of The Vehicle

SCANeR Studio simulator is the most comprehensive simulation platform addressing prototyping, validation, training of ADAS, and navigating Autonomous Vehicles as well as Human-Machine Interactions. SCANeR Studio is a driving simulation software package, where we can create a vehicle with our designed scenarios to drive the car in it (Fig. 8). Modes of driving the vehicle are: Autonomous Driving mode and Human Interaction mode.

Fig. 8.   Virtual Environment in SCANeR Studio Simulator

Autonomous Driving inputs are formulated as presented in III using a visual control algorithm using a computer vision technology which is the "Visual Servoing". Python API used to drive the vehicle, taking the camera and lidar data from the simulator.

Human Driving inputs are formulated using a joystick, it play the role of the steering wheel of a real vehicle. SIMULINK API is used to drive the vehicle using a block constructed for our joystick.

Both modes drive the vehicle in the same scenario created by the SCANeR Studio simulator.

The final command decision is given by the fusion system in the form of linear velocity and steering angle to the vehicle. Noting that, only the fusion system has the right to drive the vehicle.

V. Experimental Results

We choose to apply our methodology for an example scenario to overtake an obstacle as shown in (Fig. 8). In this case of study, the two inputs are similar since both try to avoid collision with the obstacle. While the human inputs are not admissible since the human driver was not able to avoid it correctly, but the autonomous algorithm did it successfully.
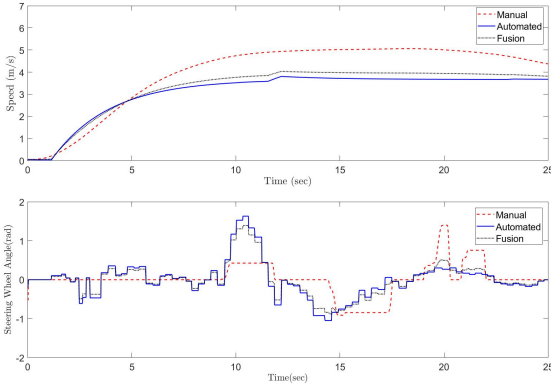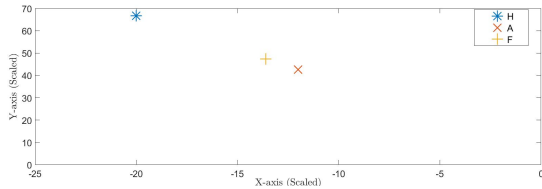

Fig. 9.   Inputs Fusion

Fig. 10.   Mapped Points 2D Presentation

As shown in (Fig. 9, 10), the fusion system sends final command input to the vehicle shifted more to the admissible input, where α is computed to a value around 0.8, thus the overtaking of the obstacle will be based on the inputs sent by the fusion algorithm which is more shifted toward the autonomous inputs ($adma > admh$).

In (Fig. 11), we were able to determine the value of the similarity and it was above the chosen threshold, thus Sim=1.

VI. CONCLUSIONS

To conclude, we were able in this work to fuse between two inputs, one from humans formulated using a joystick, and the other from an autonomous controller using a computer vision algorithm which is the "Visual Servoing". We applied our work to a scenario trying to avoid an obstacle, one was correct and the other not, and the fusion system was able to send the final command decision shifted to the higher admissible input. Simulations were done by using SCANeR Studio Simulator and MATLAB/SIMULINK.

REFERENCES

[1] S. Singh, "Critical reasons for crashes investigated in the national motor vehicle crash causation survey," Tech. Rep., 2015.

[2] https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-"levels-of-driving-automation"-standard-for-self-driving-vehicles 1/03/2021

[3] V. Thuan Nguyen, C. Sentouh, P. Pudlo and J. -C. Popieul, "Joystick Haptic Force Feedback for Powered Wheelchair - A Model-based Shared Control Approach," 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Toronto, ON, 2020, pp. 4453-4459, doi: 10.1109/SMC42975.2020.9283235.

[4] TY - BOOK Perozzi, Gabriele Sentouh, Chouki Floris, Jerome Popieul, Jean-Christophe 2020/07/11 "On Nonlinear Control for Lane Keeping Assist System in Steer-By-Wire Road Wheeled Vehicles"

[5] M. Johns et al., "Exploring shared control in automated driving," 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 2016, pp. 91-98, doi: 10.1109/HRI.2016.7451738.

[6] Zongyao Jin, Prabhakar R. Pagilla, "Human-robot teaming with human intent prediction and shared control," Proc. SPIE 11413, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II, 114130X (21 April 2020)

[7] L. Abi, D. Jin and L. Yu, "Shared Control Strategy Based on Driver's Trajectory Following Intention," 2021 International Conference on Computer, Control and Robotics (ICCCR), Shanghai, China, 2021, pp. 98-103, doi: 10.1109/ICCCR49711.2021.9349279.

[8] S. Mosbach, M. Flad and S. Hohmann, "Cooperative longitudinal driver assistance system based on shared control," 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, 2017, pp. 1776-1781, doi: 10.1109/SMC.2017.8122873.

[9] S. C. Jugade, A. C. Victorino and V. B. Cherfaoui, "Human-Intelligent System Shared Control Strategy with Conflict Resolution," 2018 IEEE 14th International Conference on Control and Automation (ICCA), 2018, pp. 686-691, doi: 10.1109/ICCA.2018.8444218.

[10] S. C. Jugade, A. C. Victorino and V. B. Cherfaoui, "Shared Driving Control between Human and Autonomous Driving System via Conflict resolution using Non-Cooperative Game Theory," 2019 IEEE Intelligent Transportation Systems Conference (ITSC), Auckland, New Zealand, 2019, pp. 2141-2147, doi: 10.1109/ITSC.2019.8917196.

[11] Shriram Jugade, Alessandro Corrêa Victorino. Grid based Estimation of Decision Uncertainty of Autonomous Driving Systems using Belief Function theory. 15th IFAC Symposium on Control in Transportation Systems (CTS), Jun 2018, Savona, Italy. pp.261-266

[12] Watson, J. (2001). Strategy: An introduction to game theory.

[13] H. Maske, G. Chowdhary and P. Pagilla, "Intent aware shared control in off-nominal situations," 2016 IEEE 55th Conference on Decision and Control (CDC), Las Vegas, NV, USA, 2016, pp. 5171-5176, doi: 10.1109/CDC.2016.7799060

[14] Xing, C. Lv, L. Chen, H. Wang, H. Wang, D. Cao, E. Velenis, andF. Y. Wang, "Advances in Vision-Based Lane Detection: Algorithms,Integration, Assessment, and Perspectives on ACP-Based Parallel Vision,"IEEE/CAA Journal of Automatica Sinica, vol. 5, no. 3, pp. 645–661, 2018.

[15] J. Kim, J. Kim, G. J. Jang, and M. Lee, "Fast learning method forconvolutional neural networks using extreme learning machine and itsapplication to lane detection,"Neural Networks, vol. 87, pp. 109–121,2017. [Online]. Available: http://dx.doi.org/10.1016/j.neunet.2016.12.002

[16] P. L. Xingang Pan, Jianping Shi, "Spatial as deep: Spatial cnn for trafficscene understanding," February 2018.

[17] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c.Woo, "Convolutional lstm network: A machine learning approach forprecipitation nowcasting,"arXiv preprint arXiv:1506.04214, 2015.

[18] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigmfor model fitting with applications to image analysis and automatedcartography,"Communications of the ACM, vol. 24, no. 6, pp. 381–395,1981.

[19] A. Cherubini, F. Chaumette, and G. Oriolo, "An image-based visualservoing scheme for following paths with nonholonomic mobile robots,"in2008 10th International Conference on Control, Automation, Roboticsand Vision. IEEE, 2008, pp. 108–113.

[20] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approachto collision avoidance,"IEEE Robotics Automation Magazine, vol. 4,no. 1, pp. 23–33, 1997.

[21] D. A. de Lima and A. C. Victorino, "A hybrid controller for vision-based navigation of autonomous vehicles in urban environments,"IEEETransactions on Intelligent Transportation Systems, vol. 17, no. 8, pp.2310–2323, 2016.